REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
12-04-2006	Final Report	10-12-2001 - 30-09-2005

4. TITLE AND SUBTITLE
Structure and Function of a Luciferase from Gonyaulax polyedra

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-02-1-0188

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Schultz, L. Wayne

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Hauptman-Woodward Medical Research Institute, 700 Ellicott St., Buffalo, NY 14203

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research, 800 N. Quincy St., Arlington, VA 22217-5000

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Unlimited

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The structure of a single light-emitting domain from the luciferase of the marine dinoflagellate Lingulodinium Polyedrum has been solved and refined to 1.8 Å resolution. The Lingulodinium luciferase (LCF) is a 140 kDa enzyme comprised of three contiguous, repeated luciferase domains, each expressing its own luciferase activity. This LCF has no sequence or structural similarity to that of the bacterial, coelenterate or firefly enzymes. The structure reported here is that of the C-terminal domain D3. The overall structure of the enzyme places it in the family of β-barrels with a characteristic β-clam fold at the core. The structure at pH 8.0 reveals that access to the interior of the barrel, and presumably the active site, is blocked by a helix-loop-helix that rests in contact with a small N-terminal subdomain. The pH/rate profile for the D3 domain and the full-length LCF reveal that the enzyme is most active at pH 6.3 and has little activity at pH 8.0. Preliminary molecular dynamics calculations indicate that histidine residues at the interface of the helix-loop-helix and the N-terminal subdomain become charged and initiate a large molecular motion that exposes the active site. Molecular modeling of the luciferin substrate and site-directed mutagenesis in the active site implicate an aspartate residue in the catalytic mechanism.

15. SUBJECT TERMS
Luciferase, Lingulodinium polyedrum, Gonyaulax polyedra, dinoflagellate, X-ray crystallography, bioluminescence

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			L. Wayne Schultz
Unclass	Unclass	Unclass	UL	7	19b. TELEPHONE NUMBER *(Include area code)* 716-898-8640

AD NUMBER	DATE	DTIC ACCESSION NOTICE
1. REPORT IDENTIFYING INFORMATION		REQUESTER: 1. Put your mailing address on reverse of form. 2. Complete items 1 and 2 3. Attach form to reports mailed to DTIC. 4. Use unclassified information only. 5. Do not order document for 6 to 8 weeks.
A. ORIGINATING AGENCY Office of Naval Research		
B. REPORT TITLE AND/PR NUMBER Structure and Function of a Luciferase		
C. MONITOR REPORT NUMBER		
D. PREPARED UNDER CONTRACT NUMBER N00014-02-1-0188		
2. DISTRIBUTION STATEMENT DISTRIBUTION STATEMENT A Unlimited Approved for Public Release Distribution Unlimited		DTIC: 1. Assign AD Number. 2. Return to requester.

DTIC Form 50 JUL 96

PREVIOUS EDITIONS ARE OBSOLETE

GRANT: # N00014-02-1-0188

PRINCIPAL INVESTIGATOR: L. Wayne Schultz

INSTITUTION: Hauptman-Woodward Medical Research Institute, Inc.

GRANT TITLE: Structure and Function of a Luciferase from *Gonyaulax Polyedra*

AWARD PERIOD: 10 Dec 2001 – 30 Sept 2005

OBJECTIVE: The main goals of this project are to solve and interpret the three-dimensional structures of the *Gonyaulax polyedra* luciferase and the luciferin binding protein using X-ray crystallography. The structural information will be used to elucidate the mechanisms by which these proteins bind the substrate luciferin and catalyze the production of light.

APPROACH: The specific aims of this project are generally directed to answering questions of structure and function. We will use standard molecular biology techniques to create the clones of interest and will utilize primarily bacterial and insect expression systems. We will use high-throughput methods to identify those constructs expressing the most soluble and stable proteins. Rapid purification of proteins will be effected by the use of engineered purification tags. Screening of crystallization conditions will occur on a high-throughput robotics system. Finally, we will use the technique of single crystal X-ray diffraction to obtain the proposed structures. In close collaboration with Dr. J. Woodland Hastings in the Department of Molecular and Cellular Biology at Harvard University, we will biochemically characterize and test hypotheses derived from structural information.

ACCOMPLISHMENTS:

LCF domain 3 structure. The structure of LCF D3 was recently determined in our laboratory and refined to 1.8 Å resolution. The overall morphology of D3 places it in the family of 10-stranded β-barrels referred to as the β-clam, first described by Sacchettini, et al. (1989). The N-terminal 67 amino acids form a subdomain comprised of two helices (α1 and α2) and two short two-stranded antiparallel β-hairpins (β1& β2; β3 & β4) (Fig. 1), which is stabilized by a small hydrophobic core and a highly organized hydrogen bonding structure.

Following the N-terminal structure is a long linker (residues 944-1026, thus including α3 and α4), which partially surrounds the β-barrel. The chain continues into the β-barrel (residues 1027-1170), interrupted only between β5 and β6 by a helix-loop-helix motif formed by α5 & α6 located at the top of the β-barrel. This motif interacts with the N-terminal helix (α2), forming a loosely associated three-helix bundle that blocks access to a cavity within the β-barrel, the putative ligand binding site. Helices α5 and α6 are arranged antiparallel to each other and are not conformationally related to helix-loop-helix calcium binding or helix-turn-helix DNA binding motifs. Following the β-barrel,

residues 1171-1218 are ordered and pack against the outer surface of the barrel and continue to an aspartic acid residue (D1218), which forms a salt bridge with an arginine (R1098), beyond which the last 23 amino acids in the peptide chain are not visible in the electron density, though they are present in the crystal.

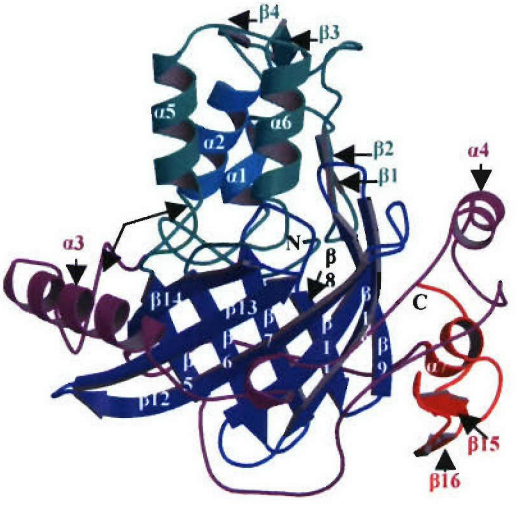

Fig. 1. Ribbon diagram of D3 LCF. The N-terminal helices and helix-loop-helix motif are at the top covering blocking access to the β-barrel.

Putative active site. The β-barrel (residues 1027-1170) of D3 is most closely related to the human fatty acid binding protein (M-FABP) with 2.1 Å RMSD for the main-chain residues comprising the β-barrel. Amino acid sequence alignment between the D3 and FABP β-barrels reveals a 19% identity and 38% similarity. It is well known that ligand binding in this family of proteins occurs within the β-barrel.

In the D3 structure at pH 8, the cavity within the β-barrel is occupied by 7 water molecules, which participate in a hydrogen-bonding network with many of the charged residues on one side of the cavity, including H1065, R1095, E1105, K1125, R1142, Q1155 and Y1168 (Fig. 2). Residues W1097 and W1117 are oriented with the indole nitrogen of the side chain pointing into the cavity and engaged in hydrogen bonds with water. Residues L1072, I1074, V1083, A1085, F1090 and F1103 form a hydrophobic patch on the other wall of the cavity. The presence of polar residues in the interior of the barrel is suggestive of an active site, as it is energetically unfavorable to bury several charged residues. While there is a significant pocket within the structure of the D3 barrel at pH 8 (volume~ 174 Å ^{3}), there is neither enough room to accommodate the tetrapyrrole (volume~ 564 Å ^{3}), as estimated by GRASP, nor an opening from the outside. A significant conformational change must thus occur to provide access to and space for a ligand in the active site.

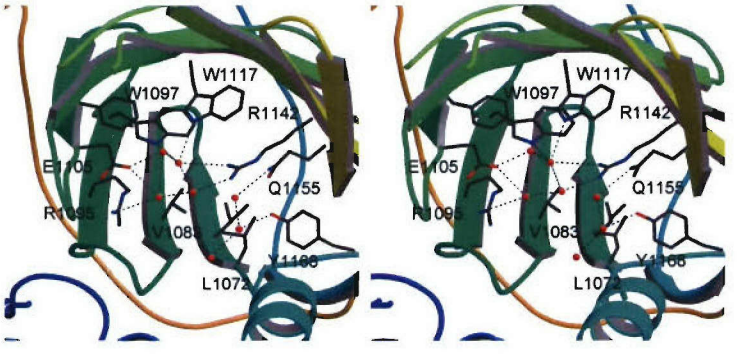

Fig. 2. Stereoview of putative active site in the β-barrel.

Site-directed mutagenesis

We have initiated site-directed mutagenesis studies to identify residues involved in catalysis of the bioluminescent reaction. Mutagenesis experiments were directed using the crystal structure (Fig 2). Table I details the progress of the mutagenesis studies.

At this point the results are qualitative as the enzymatic assays of the purified enzymes have not been initiated. Current activity is based on crude *E. coli* cell lysate (Table I). Mutations on the interior of the β-barrel appear to extinguish the ability of the enzyme to produce light. However, mutations of residues on the outside of the barrel but adjacent to active site residues do not extinguish activity. These results preliminarily identify the active site inside the barrel. Purification of the enzymes and structural work will definitively identify the active site.

Table I. Current D3 active-site mutants

Mutant	Sequence confirmed	expressed	activity	purified	Crystallization trials
K1094A	yes	yes	yes	na	na
R1095A	yes	yes	no	na	na
R1095K	yes	yes	some	yes	yes
W1097A	yes	yes	no	na	na
Y1104A	yes	yes	yes	na	na
E1105A	yes	yes	no	na	na
E1105D	yes	yes	some	yes	yes
E1105N	yes	yes	no	no	na
P1106A	yes	yes	yes	na	na
W1117A	yes	yes	no	na	na
K1125A	yes	yes	yes	na	na
W1130A	yes	yes	no	na	na
R1142A	yes	yes	no	na	na
Y1168A	yes	yes	some	na	na

na, not attempted.

Proposed mechanism for pH regulation of activity

Position of the regulatory histidines. The typical bell-shaped pH/rate profile for the full-length LCF and each of the light emitting domains is similar; both are most active at pH 6.3 and very low at pH 8.0. However, when the N-terminal 50-80 residues are removed from a domain, the relative activity at pH 8.0 is higher when compared to that at pH 6.3. Four histidine residues (pK_a ~6.5), conserved in the N-terminal region of each domain, were changed to Ala and the resultant variants studied, showing that these residues, H899, H909, H924 and H930, were indeed responsible for the pH regulation.

Our hypothesis regarding the role of these residues in the pH-controlled regulation of LCF is that they make contacts that stabilize the N-terminal domain and the helix-loop-helix. Disruption of such contacts by protonation (as in the wild-type) or mutagenesis causes the N-terminal domain and the helix-loop-helix to move away from the protein and open the catalytic active site.

Molecular dynamics suggests a mechanism for opening the active site. The fact that replacement of four histidines by alanines restores full activity at pH 8 suggests the involvement of the unprotonated histidine residues in a network of hydrogen bonds that blocks substrate access to the active site. The X-ray structure supports this hypothesis. The four histidines make contacts with and stabilize the three-helix bundle that blocks

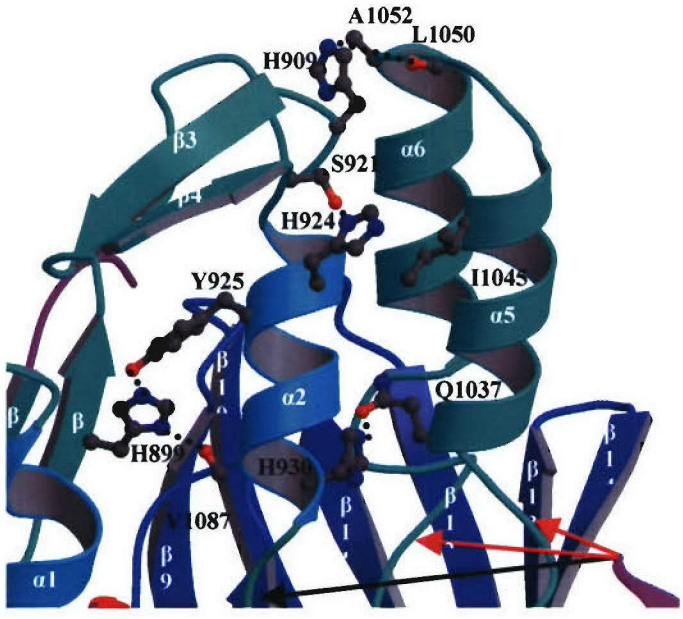

Figure 3: A detailed view of the interactions of the regulatory histidines in reference to the helix-loop-helix region. Arrows indicate the positions of the flexible Gly-Gly hinges.

entry to the putative active site. Disruption of such interaction between the helices within the bundle by protonation (as in LCF at pH 6.3), or by replacement of the four His by Ala, would be expected to cause the helices to move and open access to the catalytic site, as confirmed by molecular dynamics calculations. Starting with the X-ray coordinates but with all four His replaced by Ala, H2 and H5 are separated by 10 Å prior to simulation but by 21 Å after simulation, thus opening the solvent channel. The rest of the protein remained stable and deviated less than 1 Å from the starting model.

Three Gly-Gly sequences, one within the N-terminal helix (G933-G934) and two in the helix-loop-helix motif (G1035-1036; G1068-1069), serve as hinges about which the chains rotate (Fig. 3). Most significantly, a channel forms between the N-terminal helix α2 and the helix-loop-helix, providing solvent access to the active site. During the simulation, the active site pocket expands to 693 $Å^3$ (including the solvent channel), which is large enough to accommodate the luciferin substrate.

A model of luciferin in the active site. With the active site thus identified, a substrate fragment containing the site of oxidation (132) in the light-emitting reaction was successfully docked into the active site (Fig.4). The model places the 132 carbon near E1105 suggesting that this residue may play a role in the catalytic mechanism of light-emission.

Production and crystallization of the full-length LCF. LCF was expressed in *E. coli* as a C-terminal fusion to GST. The protein was purified using a GST affinity column and the tag removed by cleavage with TEV overnight. A second GST column was run to remove tagged protein. Following a buffer exchange to a solution containing 20mm Tris, pH 8.0, 25mm NaCl, 1mm EDTA and 1.5mm DTT, the protein was concentrated to 10mg/mL. Crystals appear from a solution of 1.5 M LiSO4 and 100mm Tris, pH 7.5 after 24 hours. The crystal morphology appears to be hexagonal plates. Diffraction studies are under way.

Since different constructs and different trials often give different crystal forms, we have also developed a His tagged system for the production of full length LCF. Quantities sufficient for crystallization have been obtained, and this will be undertaken in the near future.

Expression and purification of luciferin binding protein (LBP). A cDNA clone was provided by Dr. David Morse (U. Montréal) was inserted in-frame into T5-based

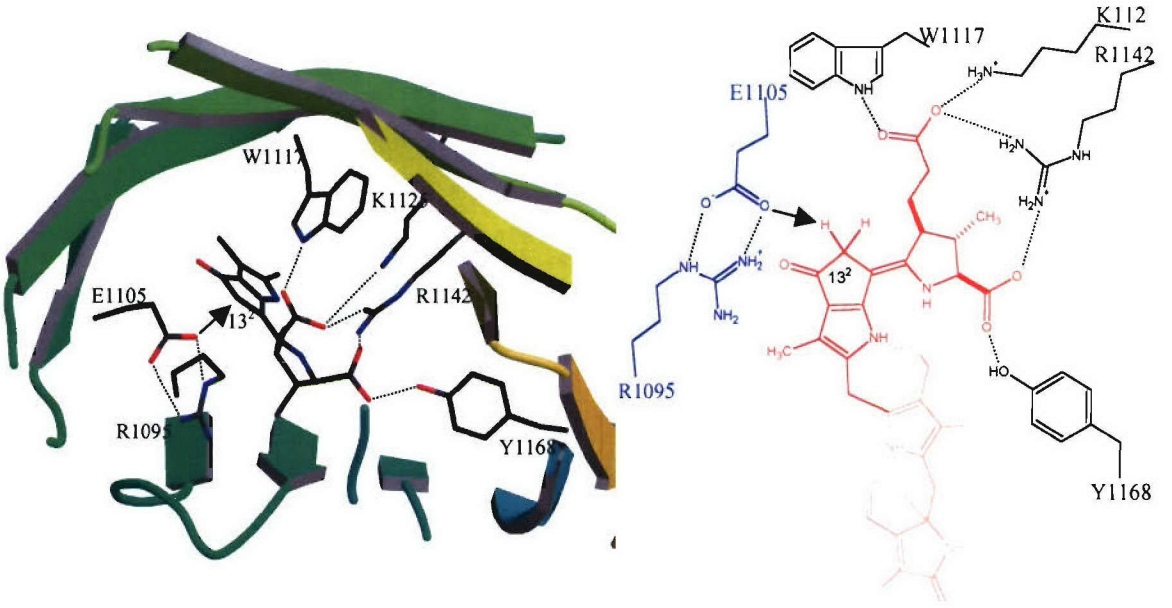

Figure 4: A model of a luciferin fragment docked into the active site of D3 LCF as defined by site-directed mutagenesis. Left, a 3D representation; right, a chemical representation.

expresson vectors (pQE30 and pQE50; Qiagen). The constructs made from pQE30 produce proteins with 6XHis N-terminal tag whereas pQE50 yield proteins with an N-terminal GST tag.

Overexpressed LBP proteins are functional, as assayed by their ability to bind the luciferin at pH greater than 7.0 and release it below pH 7.0, a property essential to the native LBP. Based on preliminary results, we selected LBP-pQE30 for the scale-up preparations, primarily because of its relatively high yields (~4 mg protein per liter of culture after affinity purification). Attempts to obtain larger amounts of LBP resulted in protein aggregation. A systematic screening of LBP solubility under various conditions indicates that a minimum ionic strength of 200 mM NaCL and at least 5% glycerol have to be maintained to keep the recombinant LBP fully soluble. Our protocol for LBP purification thus includes the elution from Ni-NTA column with the elution buffer (50 mM Tris-HCL, pH 8.0, 200 mM NaCl , 5% glycerol , 1 mM DTT and 250 mM imidazole) and fractionation on Sephacryl-100 column equilibrated with the buffer (20 mM Tris-HCL, 1 mM EDTA, 200 mMNaCl , 5% glycerol , 1 mM DTT). The LBP protein purified in this way is free of precipitate, of high purity (>97%) and stable on ice for an extended time (> 2weeks) (Fig. 5). We are now producing a large quantity of this protein for crystallization trials.

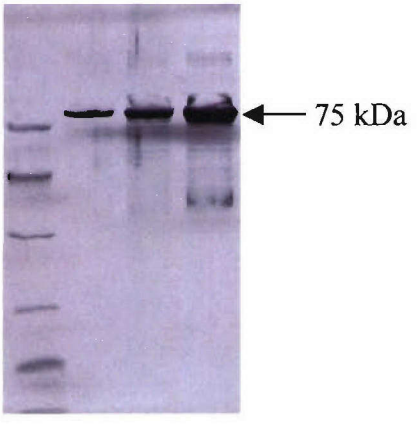

Fig 5. SDS-Page gel showing soluble purified LBP in 10, 20 and 50 ng amounts.

SIGNIFICANCE: With the structure in hand, we have identified the active site and proposed a mechanism for the light producing reaction. We will screen for inhibitors and examine structures of those inhibitors as well as the luciferin substrate bound to the active site of the enzyme. This luciferase does not appear to have structural similarity to the published structures of bacterial (Fisher et al., 1996) or firefly (Conti et al., 1996) luciferases. Each of the luciferases produce light, yet catalyze the oxidation of different substrates and require different cofactors. Our structure will help to further define the important components in a luciferase system. Interestingly, the Gonyaulax luciferase does have structural homology to the fatty acid binding protein (FBP) and retinal binding protein (RBP). Both are considered beta-clam structures and bind hydrophobic molecules deep in the protein core. However, neither FBP nor RBP have a catalytic activity. This may be the first opportunity to see how this protein fold can arrange residues to catalyze a reaction.

PATENT INFORMATION: No patents applied for.

AWARD INFORMATION: Promoted to Head, Structural Biology Laboratory

PUBLICATIONS AND ABSTRACTS:

1. Schultz, W., Cegielski, M., Liu, L. and Hastings, J. W., Structural Studies of the Catalytic Domains from a Dinoflagellate Luciferase. Abstract presented 17 July 2002, 30nd Annual Meeting of the American Society for Photobiology. p. 105.

2. Liu L., Im H., Cegielski M., LeMagueres P., Schultz L.W., Krause K.L., Hastings J.W. (2003) Characterization and Crystallization of Active Domains of a Novel Luciferase from a Marine Dinoflagellate. Acta Cryst. D *59*: 761-764.

3. Schultz, W., Cegielski, M., Liu, L. and Hastings, J. W., Structure of a Dinoflagellate Luciferase Abstract presented 28 July 2003, American Crystallographic Association 2003 Meeting. p. 149.

4. Schultz, W., Cegielski, M., Liu, L. and Hastings, J. W., Structural Studies of a Luciferase from the Marine Dinoflagellate, Lingulodinium polyedrum. Abstract presented 12 July 2004, 32nd Annual Meeting of the American Society for Photobiology. p. 9.

5. Schultz, W., Cegielski, M., Liu, L. and Hastings, J. W., Structure of a Dinoflagellate Luciferase Abstract presented 15 Aug. 2004, 18th Symposium of the Protein Society. p. 127.

6. Schultz, L. W., Cegielski, M., Liu, L. and Hastings, J. W., Crystal Structure of a pH-regulated Luciferase Catalyzing the Bioluminescent Oxidation of an Open Tetrapyrrole, Proc. Natl. Acad. Sci., **102**, 1378-1383 (2005).